

⑩ 日本国特許庁 (JP)

⑪ 特許出願公開

⑫ 公開特許公報 (A)

昭59-107397

⑮ Int. Cl.<sup>3</sup>  
G 10 L 1/00

識別記号

庁内整理番号  
R 7350-5D

⑬ 公開 昭和59年(1984)6月21日

発明の数 1  
審査請求 未請求

(全 9 頁)

⑭ 音声認識装置

⑯ 特 願 昭57-217296  
⑰ 出 願 昭57(1982)12月10日  
⑱ 発 明 者 斗谷充宏  
大阪市阿倍野区長池町22番22号  
シャープ株式会社内

⑲ 発 明 者 外川文雄  
大阪市阿倍野区長池町22番22号  
シャープ株式会社内  
⑳ 出 願 人 電子計算機基本技術研究組合  
東京都港区三田1丁目4番28号  
㉑ 代 理 人 弁理士 福士愛彦 外2名

明 細 書

1. 発明の名称

音声認識装置

2. 特許請求の範囲

1. 認識すべき音声の特徴パターンと、予め登録された複数種類の音声の特徴標準パターンとの類似度計算を行なって、認識判定を行なう認識装置において、

上記複数種類の特徴標準パターン毎にカウンタ手段を備えたことを特徴とする音声認識装置。

2. 前記カウンタ手段は認識すべき音声の認識判定結果に応じてそのカウンタ値を増減する手段を含む特許請求の範囲第1項記載の音声認識装置。

3. 前記カウンタ手段は該カウンタ値に応じて特徴標準パターンの更新を制御する手段を含む特許請求の範囲第1項記載の音声認識装置。

3. 発明の詳細な説明

<技術分野>

本発明は認識すべき音声の特徴パターンと、予

め登録された複数種類の音声の特徴標準パターンとの類似度計算を行なって認識判定を行なう音声認識装置の改良に関し、更に詳細には予め登録されている音声の特徴標準パターンの良否を知る手段を与えるようにした音声認識装置に関するものである。

<従来技術>

従来より複数の特徴標準パターンを登録しておいて、その標準パターンと入力特徴パターンとのマッチングによって音声認識を行う装置が実用化されているが、このような音声認識装置において特徴標準パターンを一度登録すると、この登録されている特徴標準パターンの良否を定量的に知ることが出来ず、認識結果から経験的に特徴標準パターン(音声パターン)の良否を判断する必要があった。

<目的>

本発明は上記の点に鑑みて成されたものであり、予め登録されている特徴標準パターンの良否を定量的に知る手段とを、またその特徴標準パターン

の良否に応じて、選択して登録パターンへの入れ換えを行うことができるようにした音声認識装置を提供することを目的としている。

#### ＜実施例＞

以下、本発明を一実施例を挙げて詳細に説明する。

第1図は本発明の一実施例装置の構成を示すブロック図であり、単語単位に発声された音声を単語単位に認識し、複数の単語候補に対して辞書照合を行い、認識結果を出力する音声認識装置を例にして示している。

第1図において1は音声入力をピックアップするマイクロホン、2は単語単位に発声され上記マイクロホン1を介して入力された音声を単語毎に分析して入力パターンとし、標準パターンメモリ3に記憶された標準パターンと入力パターンとのマッチングを行ない認識結果を出力する単語認識部、3は登録された標準(特徴)パターンを保持する標準パターンメモリ、4は上記標準パターンメモリ3に記憶された各標準パターンに対応

上記標準パターンメモリ3及びカウンタメモリ手段4への情報の初期登録動作は第4図に示す初期登録フローに従って行われる。

即ちキー入力装置7の所定キーを操作して装置を初期登録動作モードに設定すると、CPU10の制御の下に表示装置8に発声すべき単語、例えば「あ」が表示される(ステップn1, n2)。オペレータは表示装置8に表示された単語を確認して音節を発声すると、この発声された音節がマイク1を介して入力され(n3)、単語認識部2で分析されて入力音声(単語)に対する特徴パターンが作成され、この分析された入力パターン(特徴パターン)がCPU10により標準パターンメモリ3の所定位置(例えばあ<sub>A</sub>に対応した位置)に記憶されると共に(n4)、この登録された標準パターン(あ<sub>A</sub>)に対応したカウンタメモリ手段4の所定位置に初期値「80」がセットされる(n5)。このような一連の動作が標準パターンの全てに対して行なわれ、この結果カウンタメモリ手段4の各標準メモリに対するカウン

して所定のカウンタ値を記憶するカウンタメモリ手段、5は辞書照合時に必要な単語を記憶している単語辞書メモリ、6は標準パターンのテスト用の単語が複数個記憶されている標準パターンテスト用単語メモリ、7はキーボード入力装置であり、例えば第2図に示すようにかなキー7a、単語入力の終了及び次候補を呼び出すための「変換/次候補」キー7b、認識結果の確定を指示する「確定」キー7c、認識結果の修正を指示する「修正」キー7d等が備えられている。また8は認識結果等を表示する表示装置、9は標準パターン等の退避に用いられるフロッピーディスク装置、10は上記各装置2～9を制御するコントローラ(CPU)である。

上記標準パターンメモリ8には「あ」～「ん」までの単語の特徴パターンがそれぞれ5個(A～E)ずつ記憶されている。また上記の各標準パターンに対応するカウンタメモリ手段4にはそれぞれ第3図に示すように例えば初期値「80」が設定記憶される。

タ値が第3図に示すようにそれぞれ初期値「80」に設定される。

次に上記のようにしてある値(例えば「80」)に設定されたカウンタメモリ手段4の値が認識動作等に応じて増減する動作について説明する。

#### (1) 認識時のカウンタ値の増減

認識時の処理フローが第5図に示されており、入力音声「あかい」を認識する場合を例にして説明する。

今装置を認識動作モードにして認識すべき音声、例えば「／あ／／か／／い／」(赤い)を発声すると、この音声マイク1を介して入力され(n11, n12)、単語認識部2において入力音声単語ごとに順次認識され、「あ」を認識した結果として「あ<sub>B</sub>」、「は<sub>D</sub>」、「あ<sub>C</sub>」、「ば<sub>A</sub>」という順序で標準パターンに近かったことを示す認識単語候補が得られる。次に「か」が認識され、同様に「い」が認識され、その結果第1表の如き各音節の認識結果が得られる(n13)。

第1表	あ <sub>B</sub>	か <sub>C</sub>	ひ <sub>C</sub>	…第1位
	は <sub>D</sub>	あ <sub>A</sub>	い <sub>B</sub>	…第2位
	あ <sub>C</sub>	か <sub>A</sub>	き <sub>E</sub>	
	ば <sub>A</sub>	か <sub>B</sub>	い <sub>D</sub>	

ここで、単語音声入力の終了であることをキー入力装置7の「変換」キー7bで指示入力すると(n14)、CPU10の制御の下に第2表に示す如き音節ラティスが作成される。

第2表	あ	か	ひ
	は	あ	い
	ば		き

次にこの音節ラティスから単語としての候補列が、その値からしさの順で第3表の如く作成される。

第3表

いた単語であれば「確定」をキー入力装置7のキー7cを操作して、そうでなければ「次候補」をキー入力装置7のキー7bを操作して、次の認識結果を表示装置8に表示させる。また入力しようとした単語が候補になかった場合は「修正」キー7dとかなキー7aにより修正した後に「確定」キー7cを入力することになる(n17~n22)。

この例の場合には入力通りに「あかい」が第1候補になっているため、すぐに「確定」キー7cを操作することになる(n17, n18)。この操作によりステップn23、n24に移行し、まず第1音節(n=1)の認識結果を調べる。ステップn25において第1位の標準パターンが正解であるか否かが判断され、今「あ」に対する認識結果は上記第1表に示すように第1位が「あ<sub>B</sub>」の標準パターンであり、正解と判断されステップn26に移行してカウンタメモリ手段4の「あ<sub>B</sub>」の標準パターンに対応した位置の内容が「+2」される。次にステップn28に移行して第2位の標準パターンが正解であるか否かが判断され、今

①	あ	か	ひ
②	は	か	ひ
③	あ	か	い
④	あ	あ	ひ
⑤	は	か	い
⑥	は	あ	ひ
⑦	あ	あ	い
⑧	は	あ	い
⋮			
⑮	ば	あ	き

その後、作成された一つ一つの候補列と単語辞書メモリ5に記憶された単語情報との照合が行なわれ、一致した第4表の如き候補列がCPU10内の辞書照合結果バッファに記憶され(n15)、照合結果の第1候補が表示装置8に表示される。

第4表

①	あ	か	い	(赤い)
②	は	か	い	(破壊)

オペレータは表示を確認し、入力しようとして

第1表に示すように「あ」に対する第2位が「は<sub>D</sub>」であるため、ステップn30に移行してカウンタメモリ手段4の「は<sub>D</sub>」の標準パターンに対応した位相の記憶内容が「-1」され、次にステップn31に移行してn=2にカウントアップされ、n32を通過してn24に戻り、第2音節の認識結果を調べる動作に移行し、以下同様の動作が繰返される。

今第2番目の音節「か」に対する認識結果は第1位が「か<sub>C</sub>」であるため正解と判断されてカウンタメモリ手段4の「か<sub>C</sub>」に対する値が「+2」され、第2位は「あ<sub>A</sub>」であるので不正解と判断されてカウンタメモリ手段4の「あ<sub>A</sub>」に対する値が「-1」される。また第3番目の音節「い」に対する認識結果は第1位が「ひ<sub>C</sub>」であるため不正解と判断されて、カウンタメモリ手段4の「ひ<sub>C</sub>」に対する値が「-2」され、第2位が「い<sub>B</sub>」であるため正解と判断されて、カウンタメモリ手段4の「い<sub>B</sub>」に対する値が「+1」される。

以上のような動作により、標準パターンのカウントメモリ手段4の内容は第6図のように変化する。

#### (II) テストモードにおけるカウンタ値の増減

テストモードの処理フローが第7図に示されており、音声認識装置において、テストモードの動作がスタートすると(ステップn41)、標準パターンテスト用単語メモリ6に記憶された単語(発声すべき単語)が順次読み出されて表示装置8に表示されることになる(n42)。

今「あかい」という発声の指示が表示装置8上に成された場合(n42)、オペレータはその表示装置8に表示された単語を発声して音声を入力する(n43)。この入力音声「／あ／／か／／い／」は単音節認識部2において単音節ごとに順次認識され、上記した認識モードの時と同様に各音節ごとに認識結果が例えば上記した第1表に示すように得られる。

テストモードの時には、発声された単語が確定しているため、以降は認識モードにおける「確定」

CPU10に与える(n63)。

CPU10は与えられた1位の認識結果が入力パターンと一致しているか否かを判断して(n64)、一致していない場合には、何らかのハードエラーが生じたものと判断して、そのエラー表示を表示装置8にて行ない、以後の動作を中止する(n65、n66)。

また1位の認識結果と入力パターンが一致している場合にはステップn67に移行して2位～6位の認識結果をチェックすることになる。

今、例えば第5表に示すように「あA」を入力パターンとして認識処理を行った結果、1位が「あA」であり、2位以下6位までがそれぞれ「あC」、「はA」、「あE」、「かB」、「あD」であった場合、入力パターンと同一カテゴリーのものが「あC」、「あE」、「あD」の3個であると判断され(n68)、入力パターンの評価値として「3」を2倍した「6」が算出され(n69)、この値が「あA」の標準パターンに対応したカウンタメモリ位置の値に加算され(n71)、次に

キーの操作以後(n23)の処理と同じ動作(ステップn46～n57)が行なわれる。

またこれらの一連の動作は標準パターンテスト用単語メモリ6から順次単語を読み出しては表示して、同様の処理が実行される。

なお標準パターンテスト用単語メモリ6には各音節が均等に出現するように単語を選定して入れておくことにより、全ての標準パターンについてテストを行なうことが出来る。

#### (III) 標準パターン間の認識処理におけるカウンタ値の増減

標準パターン間の認識の処理フローが第8図に示されており、音声認識装置において標準パターン間の認識処理の動作がスタートすると(ステップn61)、CPU10は標準パターンメモリ8に記憶された標準パターンの一つを選び出して、単音節認識部2に入力パターンとして与える(n62)。この入力パターンに対して、単音節認識部2においてメモリ8に記憶されている各標準パターンとのマッチングを行ない、その認識結果を

ステップn72を辿ってn73に移行して次の標準パターンが選択されてステップn62に戻り同様の動作が行なわれる。

入 力 パターン	認 識 結 果						<評価点>
	1位	2位	3位	4位	5位	6位	
あA	あA	あC	はA	あE	かB	あD	6
あB	あB	はB	はC	かC	ばE	あD	2
⋮							
はA	はA	あA	あB	かB	あD	あC	-4
⋮							

第5表

上記第5表に示すように「あB」を入力パターンとした場合には、2位～6位の間に同一カテゴリーのものが一つしかないため、評価値は2となる。また「はA」の場合には、2位～6位の間に同一カテゴリーのものがいないため、評価点は「0」となるが、非常に同一カテゴリーのものから遠いことを強調する意味でステップn68からn70に移行して「-4」の評価点が与えられる。

このような動作を全標準パターンに関して行うことにより、各標準パターンの良否がカウンタメモリ手段4に記憶された値により、定量的に知ることが可能となる。

なお、上記の説明において用いた評価値は、一例に過ぎず、認識結果の順位による重み付け、6位以下の結果も用いる等の種々の変形が可能である。

上記のようにして増減されたカウンタメモリ手段4の値を用いて特徴標準パターンの更新を制御する方法について説明する。

初期の音声標準パターンの登録時にはカウンタメモリ手段4の各記憶位置には「80」がセットされ、その後の認識処理動作によってカウンタメモリ手段4の内容が変化し、例えば第9図に示すように変化したものとする。

このような状態において、不良の音声標準パターンを入れ換える更新登録モードにおいて、上記のカウンタメモリ手段4の記憶内容が参照されることになる。

ターンを最も悪いものから入れ換えられる。

以上のようにして悪い標準パターンの更新が行なわれるが、カウンタメモリ手段4の値がある程度以下になった時に、その標準パターンが非常に悪いことをオペレータに知らせ、そのパターンの更新を行なわせるようにしてもよい。

また、標準パターンチェックモードを設け、カウンタメモリ手段4の値がある程度以下の標準パターンあるいは標準パターンに対する単音節を表示に出し、それらのみの標準パターンを更新するように成してもよい。

なお上記の例では単語単位に発声された音声を単音節単位で処理を行う場合を示したが、本発明はこれに限定されるものではなく、例えば文節単位に発声された音声であってもよい。

またカウンタメモリ手段4の初期設定値を「80」としたが、この値に限定されるものではない。

更に、第1図に示した構成例においてメモリ3及び4をバックアップしていない場合等においては、上記したカウンタ値の増減あるいは標準パ

第10図に更新登録動作のフローが示されており、音声認識装置において、更新登録モードの動作がスタートすると(n81)、CPU10は最初に、発声すべき単音節を表示装置8に表示する(n82)。今「あ」が表示されたとすると、オペレータはその表示された単音節を発声して音声入力し(n83)、この音声入力された単音節(「あ」)が単音節認識部2により分析されて入力パターンが作成され、その作成された入力パターンが「あ」のカテゴリー内で最もカウンタ値の小さい標準パターン(第9図に示した例では標準パターン「あ<sub>D</sub>」)の位置に入れられ(n84)、そのカウンタの値が「80」にセットし直される(n85)。

続いて「い」が表示され、音声入力した「い」の分析結果である入力パターンが「い<sub>A</sub>」の位置に入れられ、その標準パターン「い<sub>A</sub>」に対応したカウンタメモリ手段4の位置が「80」にセットされる。

以上のような動作を繰返して各音節毎に標準パ

ターンの登録・更新を行なった場合には、標準パターンメモリとカウンタメモリ手段の内容をフロッピー・ディスク9に落としてから電源の断あるいはオペレータの交替を行ない、再使用の時にはフロッピーディスク9からその退避した内容を標準パターンメモリ及びカウンタメモリ手段にロードする必要があることは言うまでもない。

#### <効果>

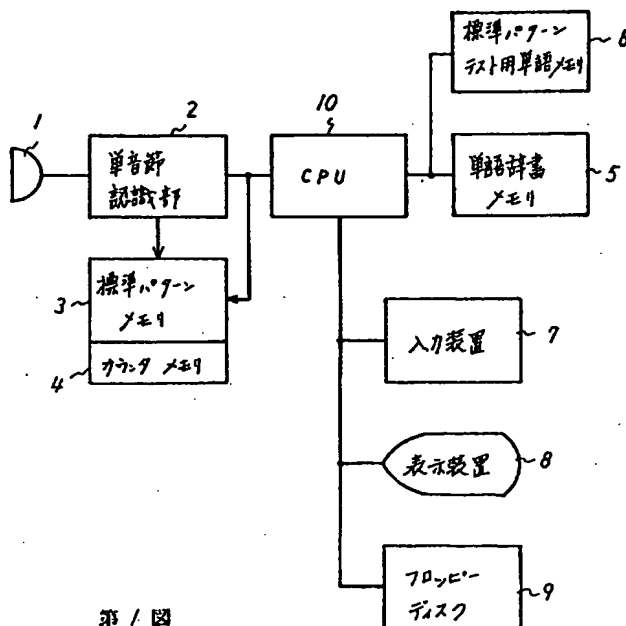
以上述べたように本発明によれば、認識すべき音声の特徴パターンと、予め登録された複数種類の、音声の特徴標準パターンとの類似度計算を行なって、認識判定を行なう認識装置において、複数種類の特徴標準パターン毎にカウンタ手段を備え、このカウンタ手段の値によって予め登録されている特徴標準パターンの良否を定量的に知ることが可能となり、この特徴標準パターンの良否に応じて選択して登録パターンを入れ換えることが出来、結果として認識率を高めることが可能となる。

#### 4. 図面の簡単な説明

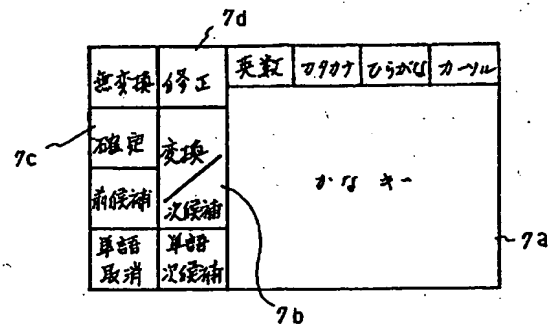
第1図は本発明の一実施例装置の構成を示すブロック図、第2図はキー入力装置の一例を示す平面図、第3図はカウンタメモリ手段の初期記憶内容を示す図、第4図は初期登録動作を示すフロー図、第5図は認識時の処理動作を示すフロー図、第6図はカウンタメモリ手段のカウンタ値の増減例を示す図、第7図はテストモードの処理動作を示すフロー図、第8図は標準パターン間の認識の処理フローを示す図、第9図はカウンタメモリ手段のカウンタ値の一例を示す図、第10図は更新登録の処理動作を示すフロー図である。

2…単音節認識部、3…標準パターンメモリ、4…カウンタメモリ手段、5…単語辞書メモリ、6…標準パターンリスト用単語メモリ、7…入力装置、8…表示装置、9…フロッピーディスク、10…CPU、11…コントローラ(CPU)。

代理人 弁理士 福 士 愛 彦 (他2名)



第1図

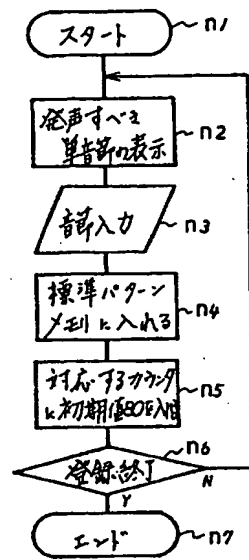


第2図

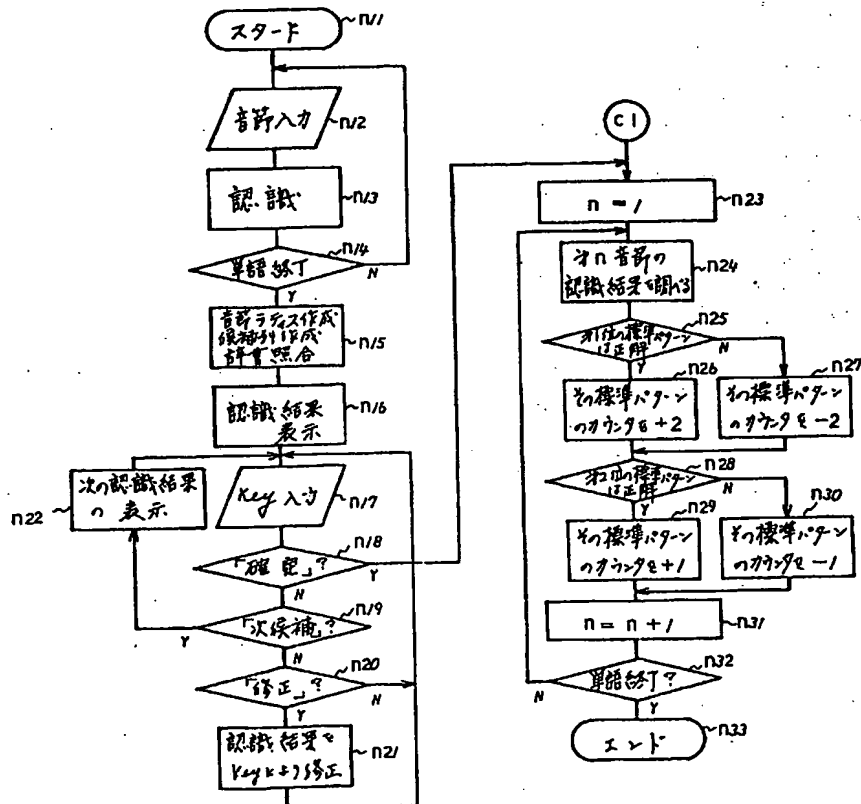
4

	あ	い	う	え	お
1	あA 80	あB 80	あC 80	あD 80	あE 80
2	いA 80	いB 80	いC 80	いD 80	いE 80
3	うA 80	うB 80	うC 80	うD 80	うE 80
4	えA 80	えB 80	えC 80	えD 80	えE 80
5	おA 80	おB 80	おC 80	おD 80	おE 80
6	かA 80	かB 80	かC 80	かD 80	かE 80
III	んA 80	んB 80	んC 80	んD 80	んE 80

第3図



第4図

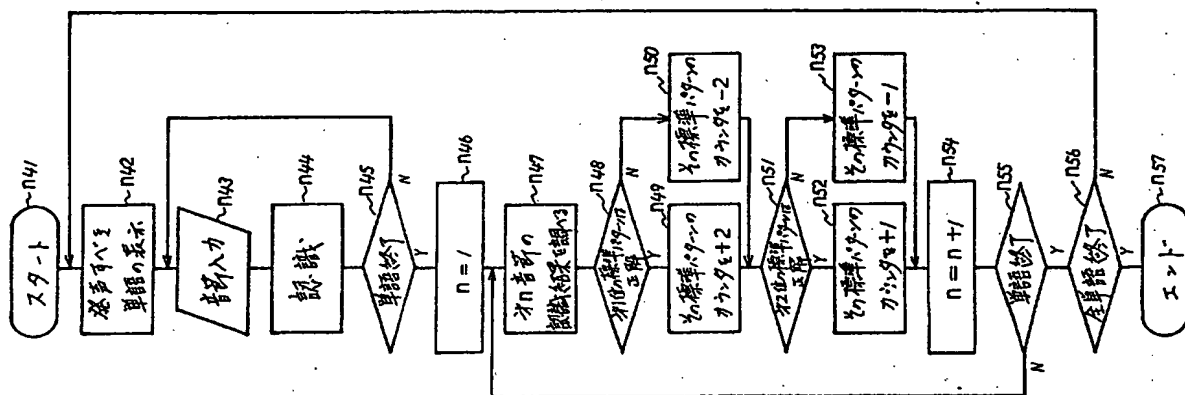


第5図

1	あ <sub>A</sub>	80	あ <sub>B</sub>	80	あ <sub>C</sub>	80	あ <sub>D</sub>	80	あ <sub>E</sub>	80
2	い <sub>A</sub>	80	い <sub>B</sub>	80	い <sub>C</sub>	80	い <sub>D</sub>	80	い <sub>E</sub>	80
3	う <sub>A</sub>	80	う <sub>B</sub>	80	う <sub>C</sub>	80	う <sub>D</sub>	80	う <sub>E</sub>	80
4	え <sub>A</sub>	80	え <sub>B</sub>	80	え <sub>C</sub>	80	え <sub>D</sub>	80	え <sub>E</sub>	80
5	お <sub>A</sub>	80	お <sub>B</sub>	80	お <sub>C</sub>	80	お <sub>D</sub>	80	お <sub>E</sub>	80
6	か <sub>A</sub>	80	か <sub>B</sub>	80	か <sub>C</sub>	80	か <sub>D</sub>	80	か <sub>E</sub>	80
III	ん <sub>A</sub>	80	ん <sub>B</sub>	80	ん <sub>C</sub>	80	ん <sub>D</sub>	80	ん <sub>E</sub>	80

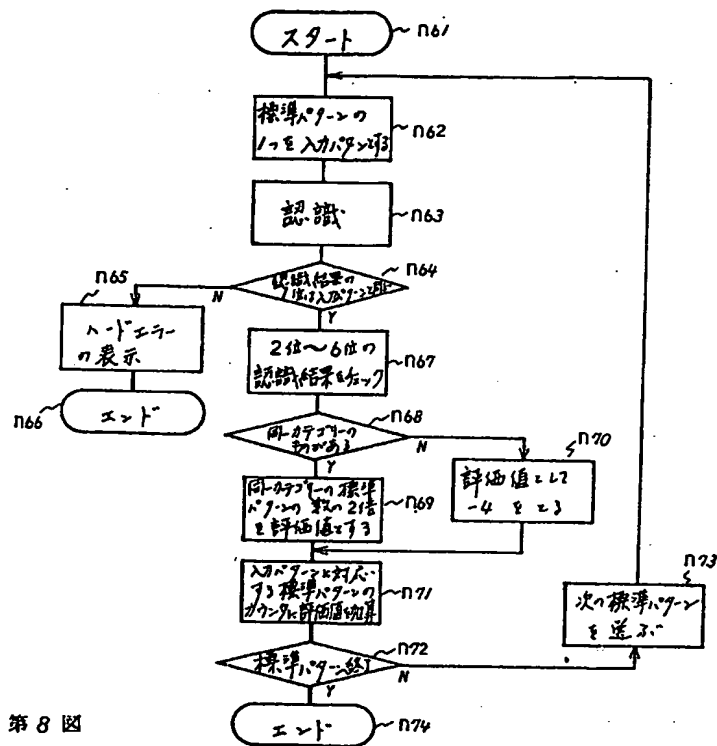
Annotations: 78 points to (あ<sub>A</sub>, 80), 82 points to (あ<sub>B</sub>, 80), 81 points to (あ<sub>C</sub>, 80), 81 points to (ん<sub>C</sub>, 80).

第6図



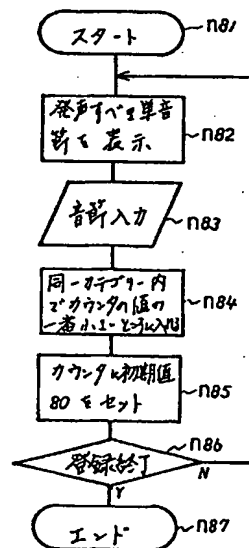
第7図





1	あ <sub>A</sub>	72	あ <sub>B</sub>	90	あ <sub>C</sub>	92	あ <sub>D</sub>	70	あ <sub>E</sub>	84
2	い <sub>A</sub>	64	い <sub>B</sub>	88	い <sub>C</sub>	82	い <sub>D</sub>	90	い <sub>E</sub>	83
3	う <sub>A</sub>	80	う <sub>B</sub>	64	う <sub>C</sub>	78	う <sub>D</sub>	82	う <sub>E</sub>	76
4	え <sub>A</sub>	98	え <sub>B</sub>	83	え <sub>C</sub>	66	え <sub>D</sub>	84	え <sub>E</sub>	62
5	お <sub>A</sub>	73	お <sub>B</sub>	78	お <sub>C</sub>	88	お <sub>D</sub>	83	お <sub>E</sub>	83
6	か <sub>A</sub>	82	か <sub>B</sub>	90	か <sub>C</sub>	75	か <sub>D</sub>	80	か <sub>E</sub>	84
110	こ <sub>A</sub>	80	こ <sub>B</sub>	80	こ <sub>C</sub>	78	こ <sub>D</sub>	75	こ <sub>E</sub>	81
111	く <sub>A</sub>	86	く <sub>B</sub>	86	く <sub>C</sub>	84	く <sub>D</sub>	80	く <sub>E</sub>	80

第9図



第10図